DOCUMENT RESUME

ED 208 565 EA 014 139

AUTHOR Holloway, William H.
TITLE A Critique of the Evaluability Assessment Model in Organizational Analysis.
PUB DATE Oct 81
NOTE 17p.

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Evaluation Criteria; Evaluation Methods; *Holistic Evaluation; *Organizational Effectiveness; Organizational Objectives; Program Descriptions; Program Effectiveness; *Program Evaluation; Public Agencies
IDENTIFIERS Evaluability Assessment; Organizational Analysis

ABSTRACT

A critique of a program evaluation technique called the evaluability assessment model (EAM), this paper describes the technique and its use in evaluating a federal agency. As defined here, EAM is divided into several steps. In the first step, the program description is reviewed to determine whether it is well-enough conceived to be a basis for adequate program performance. In this stage (the determination of "evaluability") evaluators look for such things as clarity of goals, acceptability to policy-makers, and congruence of program descriptions with program activities. If a positive determination of "evaluability" is made, then evaluators assess whether program evaluation is likely to be useful in improving program performance. If so, the best evaluation methods are chosen and the evaluation proceeds. After briefly describing the use of EAM in evaluating a federal agency, this critique notes several problems with the way the method was used, including the difficulty of validating or agreeing on such things as the mission or activities of the agency, the best way to declare an agency "nonevaluable," and how to treat the many nonevaluable agencies. In spite of such problems, however, EAM was found to permit a holistic view of organizations and to be sufficiently functional to warrant continued use.
(Author/JM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED208565

A Critique of the Evaluability Assessment Model in Organizational Analysis

William H. Holloway, Ph.D.
Associate Professor of Administration
Department of
Educational Policy and Administration
School of Education
University of Kansas

A paper prepared for presentation and publication review. As such no part of the paper may be copied without the express permission of the author.

October 1981

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

W.H. Holloway

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 014 139

A Critique of the
Evaluability Assessment Model
in Organizational Analysis

Abstract

In recent times a number of agencies and organizations have been subjected to pressure for increased accountability on the part of their governing or policy making bodies. A number of these agencies (particularly federal) have utilized the Evaluability Assessment Model (EAM) as the methodological guideline for organizational analysis leading toward increased accountability in the sense that systematic and comprehensive knowledge of the organization is acquired and communicated.

The report is a critique of the application of EAM in one such organization where the work was completed by an external evaluation team. Among the problems noted in the effort and further discussed in the paper are (1) the validation of logic, function, and measurement models, (2) the declaration and treatment of a "non-evaluable agency", (3) address of the problem of real-ideal discrepancies, and (4) heirarchical ordering and integrating of function and logic displays. Conclusions reached on the basis of the experience include (1) the systems-oriented focus of EAM permits a multi-level, wholistic view of organizational activities and achievements, (2) the model is qualitatively oriented and independent of organizational complexity, (3) the model is sufficiently functional to warrant continued use and (4) the noted problems deserve consideration prior to future use which should strengthen the results achieved.

Special praise and recognition are due several individuals of the evaluation task team without whose ideas, interactions and other contributions this effort could not have succeeded. In particular, credit is due Dr. Bill Burgess, Project Director, and Dr. Kathryn Stiano, Project Coordinator.

A CRITIQUE OF THE EVALUABILITY ASSESSMENT MODEL IN ORGANIZATIONAL ANALYSIS

William H. Holloway, Ph.D.
University of Kansas

Whatever else research on organizations has produced or affirmed, the composite findings would not argue against the assertion that organizations are dynamic, evolving entities. This attribute, perhaps more than any other, has rendered a substantial part of past research meaningless and necessitated considerable change in contemporary research methodology. Organizations do change, such changes can be constructive or disfunctional, and there is precious little information to guide researchers or administrators regarding the critical factors or parameters of change. A case in point is that less than ten years ago organizations were pressed to be specific about goals, enabling and performance objectives and the strategies that would assure accomplishment. Now the literature addresses the phenomena of "loosely coupled" organizations and agreement between what is planned and what is implemented may be an unnatural expectation.

Organizations continue to invest substantially in evaluation. Numerous accounts attribute this phenomenon to a continuing press on propensity for accountability. However, on this issue the determinants are less clear. What is clear is that evaluations of program, professional staff, service delivery, and performance continue unabated at no little cost, using a variety of methodological approaches, and producing results and a use of the results which are at least suspect if not an outright condemnation. Questions arise as to whether we have failed to account for the dynamic attributes of organizations, whether the evaluation models are not yet capable of producing valid comprehensive conceptions, or whether the decision makers are simply ill-

prepared or unwilling to act on the information provided?

This report is a critique of the use of a new evaluation model, evaluability assessment (EA), in the organizational analysis of a federal agency. Evaluation models range from the early Tylerian focus on objectives to goal-free evaluation and from simple (the use of selected parts of a model) to comprehensive as in the case of the Stufflebeam CIPP model or in this case, Evaluability Assessment. Hence, evaluability assessment was designed to be comprehensive. Furthermore, it was designed to produce a clear conception of the organization and to insure that decision makers would take appropriate action. Finally, whether valid or otherwise, the model assumes that organizations are tightly coupled and highly structured or will be at the conclusion of the evaluation effort.

Background to the Study

The study involves three components for which some background information may be useful. The first is a description of the agency which was the subject of the external evaluation, the second is the particular evaluation method employed, and the third the implementation experience.

The XYZ Agency

The XYZ agency is one of many in the Department of Health and Human Services (HHS) and the Office of Human Development Services (OHDS). The various agencies within OHDS are charged with the administration of specific human service programs (aging for instance) though XYZ services an identified population and is an exception for that reason. As a result, the program effort of XYZ is less clearly defined by virtue of title and, therefore, more dependent on the 1974 Act[1] which authorizes

[1] P.L. 93-644, The Native American Programs Act (as amended in 1978 by P.L. 95-568), 1974.

the program. Most of the activity of the agency relates to (1) financial assistance through grants, (2) providing technical assistance and training and (3) the contractual conduct of research, demonstration and pilot projects all focused on the promotion of increased self-sufficiency among these people. Finally, the XYZ agency is the only organization which is authorized to serve all Native Americans without regard to place of residence, tribal or group affiliation including terminated tribes, non-federally recognized tribes, state-recognized tribes, rural non-reservation tribes, and urban groups.

The XYZ agency is headed by a Commissioner with a staff of 47. Seven of these are regional representatives housed outside the agency in offices across the nation. The remainder include several unit administrators and a large number of program specialists responsible for grant administration. The annual budget of the agency (1981) was 33.8 million dollars (reduced to 28.0 for 1982) and reveals only nominal growth from the 30.9 million dollars authorized in 1972. In contrast, the caseload of the agency (number of grantees served) has grown from 94 in 1972 to more than double (200) projected for 1982. Clearly, the growth in number of grantees served by XYZ has outdistanced the agency's funding capability. As a result, the level of funding available to each has diminished by more than half.

The operational philosophy of XYZ suggests a belief that self-sufficiency is best nurtured by locally developed strategies designed to address perceived local needs. In the case of federally recognized communities, the agency supports the mandates of P.L. 93-638, Indian Self-Determination, and the strengthening of tribal governance.

However, the agency also acts to facilitate the development of organizational capacity on the part of non-sovereign agencies.

The Evaluability Assessment Model

The United States Congress (among many others) continues to press federal agencies for accountability. Initially, the agencies treated the demand as a "pass-along" obligation and grantees throughout the nation were required to include evaluation as a specific project component. More recently, however, the Congress has clarified its expectations and mandated good planning, execution, and evaluation by the federal agencies. In the latter function, the intent was clarified to include evaluation of grantees as well as the agencies themselves. Consequently over the past four or five years, many such evaluations have occurred and often the choice of model was Evaluability Assessment (Wholey, 1979).

The design parameters for the EA model were chosen to address several of the prevalent and persistent problems of past evaluation strategies. That is, EA was designed (1) to provide an alternative procedure for agencies termed "unevaluable", (2) to generate highly specific diagrams (models) of agency activities (function models), functional plausibility (logic models), and assessment processes (measurement models), (3) to encourage frequent interaction with managers and to communicate intermittent results, (4) to involve the total heirarchy from policy maker to grant recipient, and above all (5) to produce results both in terms of a performance assessment as well as information that is received and acted upon by managers.

One perspective on the EA model can be achieved by an examination of the criteria by which an "evaluable" organization is determined

(Schmidt, et.al., 1978). All successive stages of the process involve systematic analysis of the program description to determine:

1. It is complete: the extent to which program objectives are well defined and the extent to which program performance indicated measures are clearly delineated;
2. It is acceptable to policy makers: the extent to which managers and policy makers at higher levels agree on the program objectives and on the evidence required to assess the program's accomplishments;
3. It is a valid representation of the program as it actually exists: the objectives and activities actually underway in the field (at regional, state, or local levels) match management's description and intentions;
4. It is plausible: management's expectations for the program and the causal assumptions linking program activities to results are plausible; that is, it appears likely that assumptions are valid and that expectations can be accomplished;
5. The evidence required by management is realistic, reliable, and feasible to collect: the extent to which program performance information can be reliably and feasibly collected and can realistically assess progress toward objectives;
6. The uses of evaluation information are related to the improvement of program performance: the extent to which management's intended uses of evaluation information can realistically be expected to affect program management and performance.

The determination that an agency is "unevaluable" requires an explanation which Wholey (1979) provides.

> "Evaluability assessment determines not whether the program can be evaluated (every program can be evaluated) but whether program performance is likely to be adequate and whether program evaluation is likely to be useful in improving program performance." (p. 18)

Whether or not the performance of a program is likely to be adequate constitutes the plausibility decision. That is, is it reasonable to expect that given the available resources, the specified activities, and the underlying assumptions--stated objectives will be achieved at some level and identified measures will, in fact, discern the extent of performance. For those programs (or subparts thereof) where the foregoing is true--the programs are ready to be evaluated. Failure to meet the criterion of plausibility indicates one or more of the first four criteria have not been met and steps must be taken to address noted deficiencies or performance assessment cannot succeed and appropriate decision information cannot be derived. As noted previously, these were held to be primary weaknesses in past evaluation efforts and under EA should result in a decision of unevaluable.

The latter two criteria (nos. 5 and 6) address the concurrence of managers in the choice and value of performance measures and their willingness and readiness to act on information so derived. Failure to meet these criteria also renders the program unevaluable as the latter part of Wholey's definition indicates.

In summary, EA establishes six essential criteria for program evaluability. In the event that all criteria are met management is presented with a range of options on evaluation design recognizing that information is purchased at a cost and it may not be economically

feasible to purchase a complete picture of outcomes. The options address varying subsets of measures and methods of collection as well as such strategies as comparative analysis, longitudinal, etc. Thereafter, evaluation proceeds as we commonly know it.

In the event that any criteria are not met, management is again presented a variety of options which describe actions to be taken, for example, to achieve program plausibility. Achieving plausibility also entails certain costs but is justified on the basis that the probability of accomplishment is higher than the continued evaluation of an ill-conceived program effort.

The Implementation Experience

Over a period of several months in which a team of evaluators conducted on-site visits, interviews (semi-structured), and a thorough analysis of pertinent documents, a disjoint but comprehensive picture of the agency began to emerge. The picture was fragmentary because different individuals performed separate tasks, prepared written briefs on their respective findings, and were only partially successful integrating the findings of others as shared at task force meetings. However, as is often the case, consensus did occur on various issues at intermittent points in the effort and a validated picture began to emerge.

Rather early, for example, it became clear that the respective views of policy makers, upper level managers, and agency administrators were discrepant from the U.S. Congress right on down the heirarchy. At the highest levels, the agency was almost non-existent. Formulation of the enabling legislation was completed with the assistance of lobbyists and government experts and not until refunding time did the

Office of Management and Budget (OMB) or the Congress request a report on achievements. To some extent this is understandable but overwhelming. That is, the agency's budget of 30 million dollars was only a small part of a total budget of over 3 billion dollars and the agencies request and justification occupied one-half page in a document of over 1500 pages. Clearly, the distance between policy making and implementation and the enormous disparity in relative size contributed to the reality of a loosely-coupled organization (Clark, et.al., 1981).

Also at a fairly early stage of the effort, it was established that the program of the agency would fail to meet the first criteria for evaluability and very likely all others. That is, the program was not clearly defined in the sense that goals, objectives, activities, resources, and underlying assumptions about causal linkages were missing almost without exception. In fact, though the agency's mission (self-sufficiency) was clearly understood and while both documents and verbal responses attested to congruence on goals (economic and social development), there was almost no correlation on specific objectives or the means of accomplishment (resources and activities). Hence, the assessment effort began to deal with the political ramifications of a declaration of unevaluability and the further tasks of assisting the agency to become evaluable. The course that was taken was to deal with each of the criteria in sequential order insofar as contractual responsibilities would permit. To date, the effort is rapidly closing on the completion of the fourth criteria (plausibility) and some but not all of the parts of the program are ready for evaluation.

Discussion

The problems encountered in the use of EA range from nearly insignificant to large but not disfunctional. Attention will be focused on a number of middle-range problems and only brief consideration will be given the others.

Model Validation

The Whorfian Hypothesis suggests there will be as many distinct views of a given reality as there are viewers. Other writers refer to the term "multiple realities" suggesting that each version is a true account. Thus, model validation is not a simple problem.

Evaluability Assessment requires the evaluator to formulate at least three highly specific models (function, logic, and measurement) and possibly a number of each if the organization is complex and varying levels of the program's heirarchy are needed. In addition, the substantive nature of the models must be synthesized from views gleaned from interviews, document analysis, and observation. Added to the variance among perceptions, was the tendency of respondents to vacillate between the real dimension ("what is") and the ideal dimension ("what ought to be") often without a clear clue to the shift. These circumstances render the task of achieving valid models most difficult if not impossible. The strategy employed to address this problem was to first seek consensus within the members of the evaluation team and then to reach agreement on the views with the agency's task group. The potential perils are easy to envision including the possibility of not being able to achieve consensus or agency agreement. The latter circumstance was especially true as the team attempted to achieve closure on the mission, goals, objectives, and activities of the agency.

Unevaluable Declarations

A second problem concerns the declaration that an agency is unevaluable. The problem stems not from the utility of the six criteria in reaching a determination but from a conviction that few, if any, organizations can successfully meet all six requirements of organizational programs. Thus, EA (much as the label implies) becomes a preparatory or preliminary exercise similar to an externally conducted self-study after which a variety of evaluation strategies may be purchased and implemented.

Criticism, however, is not directed toward this aspect of EA which is generally viewed as commendable but toward the failures of the authors to elaborate this possibility, spend considerably more time on contingency strategies, and in so doing, to remove some amount of the stigma that emerges from a failure to give the matter appropriate attention. Our experience revealed that certain expectations had grown on the part of the agency and the evaluation team regarding the way the task would proceed and the growing realization that the agency could not meet even the first criteria called for an abrupt and generally unanticipated change.

Real/Ideal Discrepancies

The problem associated with real/ideal discrepancies has three facets. The first facet has already been discussed in terms of the problems associated with interview data and the tendency of respondents to vacillate without visible clue. A second aspect involves the definition of respective models and the third the matter of organizational coupling.

Discrepancy analysis has long taken advantage of difference between "what is" and "what ought to be". A superficial consideration of the function and logic models would lead one to believe a parallel relationship exists. However, the definition of each clearly reveals differences (albeit small differences) and the evaluator may either act in error or as we found in our experiences, be left wondering where real and ideal fit into the overall scheme.

Function models tend to be more closely related to "what is" than logic models are related to "what ought to be". In fact, logic models attempt to depict causal (if-then) relations and underlying assumptions while function models attempt to depict "what is" if conceived prior to EA and "what ought to be" if formulated after EA. In rare cases, both instances (pre and post EA) would be the same and then the ideal would prevail in the function models.

The third facet of the real/ideal problem relates to the organizational coupling phenomena (Clark, et. al., 1981). There is increasing empirical support for the assertion that public schools, public organizations, and complex organizations are loosely coupled. That is, the design or planning function considered an expression of "what ought to be" differs considerably from the implementation function considered to be the "what is" dimension.

If this is accurate, these efforts to implement EA in public or complex organizations given that EA requires an effort to achieve tighter coupling may be appealing from an ideal (valued) perspective but disfunctional or ineffective from a practical standpoint. At best, organizations may achieve a degree of tighter coupling for a

time only to drift back to a loosely coupled state shortly thereafter. On this matter, our collective experience cannot offer assistance.

The Paradox of Models

A final problem relates to what has become known as Bonini's Paradox. Process analysis and the depiction of these findings in specific flow models is an admirable pursuit only recently come of age. However, there is a tendency to include an ever increasing amount of detail (particularly in complex organizations where heirarchical relations must also be considered) and thus to produce models so complex they become incomprehensible. At what point does the increase in detail begin to lose utility in terms of diminished understanding? There are no guidelines or easy answers. The recommendation is to be constantly alert to the potential problem and to engage in considerable testing among both evaluators and the assigned organizational task group.

Conclusions

The experience to date with the implemention of EA has led to a number of problems which are being addressed as discussed. Similarly, a number of tentative conclusions can be advanced. First, the systems perspective of the EA model permits a Wholistic view of the program of the organization which appears to be independent of complexity. Second, the model uses qualitative methods in the early stages and provides for either qualitative or quantative methods in the evaluation that is recommended after "evaluability" has been achieved. Third, though a number of problems have emerged the model is still considered sufficiently functional and a clear step forward in terms of the choices of other models as to warrant continued use.

In particular, it is believed that the model will be useful with regard to public schools though additional experiences may prove otherwise. Finally, it is concluded that the noted problems do have solutions that awareness will partially solve and further experience will completely resolve.

In summary, the EA model has much to recommend continued use and much less to argue to the contrary.

References

Clark, David L., McKibbin, Sue, and Malkas, Mark (Eds.). Alternative Perspectives For Viewing Educational Organizations. Far West Laboratory, San Francisco, California, 1981.

Schmidt, Richard, Scanlon, J. W., and Bell, J. B. Evaluability Assessment: Making Public Programs Work Better. The Urban Institute, Washington, D.C., 1978.

Wholey, Joseph S., et. al. Federal Evaluation Policy, The Urban Institute, Washington, D.C., 1970.

Wholey, Joseph S. Evaluation: Promise and Performance. The Urban Institute, Washington, D.C., 1979.